3. *On a Theorem of Weyl-von Neumann*

By Shige Toshi KURODA
Department of Physics, University of Tokyo

(Comm. by Z. SUETUNA, M.J.A., Jan. 13, 1958)

1. Introduction. Concerning the perturbation of the spectrum of a self-adjoint operator, the well-known theorem of Weyl-von Neumann[1)] states that *any self-adjoint operator in a separable Hilbert space can be changed into one with a pure point spectrum by the addition of a suitable completely continuous, self-adjoint operator with arbitrarily small Schmidt norm.*

This theorem is no longer true if "Schmidt norm" is replaced by "trace norm". This is a direct consequence of a recent result of Kato,[2)] according to which *the absolutely continuous part of the spectrum of a self-adjoint operator is never changed by the addition of a self-adjoint operator with finite trace norm.*

There still remains a gap between these two results, for there are plenty of classes of completely continuous operators other than the Schmidt class and the trace class. These classes are most conveniently described in terms of the *cross norm* introduced by von Neumann and Schatten.[3)] The purpose of the present note is to fill in the gap by showing that the theorem of Weyl-von Neumann is true for all unitarily invariant cross norms with the single exception of the trace norm (or its equivalent). This shows at the same time that the trace class is the only allowable class in Kato's theorem, as long as we are concerned with classes defined in terms of unitarily invariant cross norms.

We give a brief exposition of the properties of cross norms needed in the sequel.[4)] Let $\mathfrak{H}$ be a separable Hilbert space, $\mathbf{B}$ the space of all bounded linear operators on $\mathfrak{H}$ to $\mathfrak{H}$, $\mathbf{S}\subset\mathbf{B}$ the Schmidt class, $\mathbf{T}\subset\mathbf{S}$ the trace class and $\mathbf{F}\subset\mathbf{T}$ the space of all operators of finite rank. We denote by $\| \ \|$ the ordinary norm, by $\| \ \|_2$ the Schmidt norm and by $\| \ \|_1$ the trace norm. In conformity with Schatten's terminology,[3)] a norm $\alpha(X)$ defined on $\mathbf{F}$ will be called a *unitarily*

1) J. von Neumann: Charakterisierung des Spektrums eines Integraloperators, Actualités Sci. Ind., **229**, Paris (1935).

2) T. Kato: Perturbation of continuous spectra by trace class operators, Proc. Japan Acad., **33**, 260-264 (1957). Here the result is obtained in a general, not necessarily separable, Hilbert space.

3) R. Schatten: A theory of cross spaces, Ann. Math. Studies, Princeton (1950).

4) Detailed results concerning unitarily invariant cross norms can be found in Schatten's work cited above.

invariant cross norm, if it satisfies the following two conditions.

1) $\alpha(UXV)=\alpha(X)$, for any pair of unitary operators U, V and for any $X\in\mathrm{F}$.

2) $\alpha(X)=1$, for any projection X of rank 1.[5]

By virtue of the condition 1), $\alpha(X)$ depends only on the eigenvalues of $|X|=(X^*X)^{1/2}$. Hence $\alpha(X)$ can be expressed in the following form.[6]

$$(1)\qquad \alpha(X)=\alpha(X^*)=\alpha(|X|)=\alpha\left(\sum_{i=1}^{n}x_iP_i\right),$$

where n is the rank of X, $\{x_1,\cdots,x_n\}$ is the set of all positive eigenvalues of $|X|$ (degenerate eigenvalues being repeated) and $\{P_1\cdots,P_n\}$ is an arbitrary set of mutually orthogonal projections of rank 1. The following inequalities are direct consequences of the definition and (1).

$$(2)\qquad \alpha\left(\sum_{i=1}^{n}x_iP_i\right)\leq\alpha\left(\sum_{i=1}^{n}y_iP_i\right),\quad \text{if}\quad 0\leq x_i\leq y_i.$$

$$(3)\qquad \|X\|\leq\alpha(X)\leq\|X\|_1=x_1+\cdots+x_n.$$

We denote by C_α the completion of F with respect to the norm α. From the inequality (3) it follows that any Cauchy sequence with respect to the norm α forms also a Cauchy sequence with respect to the ordinary norm. Hence, for any element of C_α, there corresponds an completely continuous operator in **B**. But the general theory of cross norms shows that this correspondence is actually one-to-one.[7] Thus we may regard C_α as a subspace (in general dependent on the norm α) of **B**, consisting solely of completely continuous operators. The norm on $\mathrm{C}_\alpha\subset\mathbf{B}$ as a completion of F will be denoted also by $\alpha(\ )$.

By virtue of (3), the unitarily invariant cross norm α is *equivalent to the trace norm* on F, if and only if there exists a constant $c>0$ such that

$$(4)\qquad \alpha(X)\geq c\|X\|_1,\quad \text{for any}\quad X\in\mathrm{F}.$$

By means of these notions we can now formulate our result in the following form.

THEOREM. *Let α be a unitarily invariant cross norm which is defined on* F *and not equivalent to the trace norm. Then for any self-adjoint operator H in $\mathfrak{H}$ and for any positive number ε, there exists a self-adjoint operator $X\in\mathrm{C}_\alpha$ with the following properties.* 1) $\alpha(X)\leq\varepsilon$; 2) *the self-adjoint operator $H+X$ has a pure point spectrum.*

5) For our purpose the condition 2) will not play any essential rôle. It will serve only as a convenient normalization condition.

6) $\alpha(X)$ can be also expressed in terms of a "symmetric gauge function" Φ introduced by von Neumann and Schatten. It is defined by $\Phi(x_1,\cdots,x_n,0,\cdots)=\alpha(\sum^n x_iP_i)$. For our purpose there is no essential difference between these two expressions and in the sequel we shall use (1) for convenience.

7) See, for example, Schatten: Loc. cit., Theorem 5.6, p. 109.

Remark 1. Evidently $\| \|$, $\| \|_2$ and $\| \|_1$, each of them restricted on F, are unitarily invariant cross norms. Since F is dense in S with respect to the Schmidt norm and S is complete, $\mathbf{C}_\alpha = \mathbf{S}$ when α coincides with $\| \|_2$. Thus our theorem is an extension of Weyl-von Neumann's theorem. There are various unitarily invariant cross norms, for example the p-norm $\|X\|_p = \| |X|^{p/2} \|_2^{2/p}$ $(p \geq 1)$.

Remark 2. Similarly we have $\mathbf{C}_\alpha = \mathbf{T}$ when α coincides with $\| \|_1$. From this and the inequality (3) it follows that in general $\mathbf{C}_\alpha \supset \mathbf{T}$. In case when α is not equivalent to the trace norm, $\mathbf{T}$ is a proper subset of $\mathbf{C}_\alpha$. (For, $\mathbf{C}_\alpha = \mathbf{T}$ would imply that the two norms are equivalent, according to Banach's theorem and the inequality (3).) In this case the topology induced by the norm α in $\mathbf{T}$ is strictly weaker than that induced by the trace norm. Thus the above-mentioned assumption of Kato[2)] for the stability of the absolutely continuous spectrum is the weakest one in the sense described above.[8)]

2. Proof of the theorem

LEMMA 1. *Let α be a unitarily invariant cross norm and let $\{P_i\}$ be a sequence of mutually orthogonal projections of rank* 1. *Then,*

1) $\alpha\left(\sum_{i=1}^{n} P_i\right)\Big/n$, $n=1,2,\cdots$, *is a non-increasing sequence.*

2) *The norm α is equivalent to the trace norm, if and only if*

$$\lim_{n\to\infty} \alpha\left(\sum_{i=1}^{n} P_i\right)\Big/n > 0. \tag{5}$$

PROOF. By applying (1) to the operator $X=\sum_{i=1}^{n} P_i$ and taking $\{P_1,\cdots,P_{k-1}, P_{k+1},\cdots,P_{n+1}\}$ as $\{P_1,\cdots,P_n\}$ of (1), we have

$$\alpha(P_1+\cdots+P_{k-1}+P_{k+1}+\cdots+P_{n+1}) = \alpha(P_1+\cdots+P_n),$$

$k=1,\cdots,n+1$. Hence, by making use of the triangle inequality, we obtain

$$\begin{aligned} n\alpha(P_1+\cdots+P_{n+1}) &= \alpha(nP_1+\cdots+nP_{n+1}) \\ &\leq \alpha(P_1+\cdots+P_n)+\alpha(P_1+\cdots+P_{n-1}+P_{n+1})+\cdots \\ &\qquad +\alpha(P_2+\cdots+P_{n+1}) \\ &= (n+1)\alpha(P_1+\cdots+P_n), \end{aligned}$$

which proves 1). (4) implies that $\alpha\left(\sum_{i=1}^{n} P_i\right)\Big/n \geq c\left\|\sum_{i=1}^{n} P_i\right\|_1\Big/n = c > 0$. This proves the necessity of the condition (5). Conversely, suppose that (5) is true. For any $X \in \mathrm{F}$, let $\{x_1,\cdots,x_n\}$ be the positive eigenvalues of $|X|$. Then (1) gives

$$\begin{aligned} \alpha(X) = \alpha(x_k P_1 + \cdots + x_n P_{n-k+1} + x_1 P_{n-k+2} + \cdots \\ + x_{k-1} P_n),\ k=1,\cdots,n. \end{aligned} \tag{6}$$

8) However it can be weakened in another direction. Roughly speaking, Kato's theorem holds if, for example, $H>0$ and $|X|^{1/2}H^{-1/2} \in \mathrm{S}$. These results will be published elsewhere.

By virtue of the triangle inequality, (5) and (6) imply the existence of a positive constant c such that

$$\begin{aligned} n\alpha(X) &= \alpha(x_1P_1+\cdots+x_nP_n)+\alpha(x_2P_1+\cdots+x_nP_{n-1}+x_1P_n) \\ &\quad +\cdots+\alpha(x_nP_1+x_1P_2+\cdots+x_{n-1}P_n) \\ &\geq \alpha((x_1+\cdots+x_n)(P_1+\cdots+P_n)) \\ &\geq nc\,\|X\|_1. \end{aligned}$$

This proves the sufficiency of the condition (5).

LEMMA 2. *Let the unitarily invariant cross norm α be not equivalent to the trace norm. Then there exists a sequence of positive numbers ε_n such that $\varepsilon_n \downarrow 0$, $n\to\infty$, and*

$$\alpha(X)\leq \varepsilon_n n\,\|X\|, \textit{ for any } X\in \mathrm{F} \textit{ of at most rank } n.$$

PROOF. X is of rank $r\leq n$. Let $x_1,\cdots,x_r$ be the positive eigenvalues of $|X|$ and let $x_{r+1}=\cdots=x_n=0$. From (1) and (2) it follows that $\alpha(X)=\alpha\left(\sum_{i=1}^{r}x_iP_i\right)=\alpha\left(\sum_{i=1}^{n}x_iP_i\right)\leq(\max_{1\leq i\leq n} x_i)\alpha\left(\sum_{i=1}^{n}P_i\right)=\|X\|\,\alpha\left(\sum_{i=1}^{n}P_i\right)$. Hence the lemma is a direct consequence of Lemma 1.

Proof of the theorem. The proof can be carried out analogously to the proof given by von Neumann in the case of the Schmidt norm. In order to manifest the point which requires modification, we shall partly restate his proof in slightly modified form. Let $H=\int_{-\infty}^{+\infty}\lambda\,dE(\lambda)$ be a self-adjoint operator, l a fixed positive number, n a positive integer and f a fixed element of $\mathfrak{H}$. We set

$$E_m=E((2m-n)l/n)-E((2m-n-2)l/n),\ m=1,\cdots,n$$
$$E_m\mathfrak{H}=\mathfrak{M}_m,\ g_m=E_mf,\ \varphi_m=g_m/\|g_m\|\in\mathfrak{M}_m.$$

(Set $\varphi_m=0$ if $g_m=0$.) We denote by P the projection on the closed subspace determined by $\varphi_1,\cdots,\varphi_n$. As was shown by von Neumann, we obtain after a simple calculation,

$$\|(1-P)H\varphi_m\|^2\leq l^2/n^2. \tag{7}$$

From this von Neumann deduced that $\|(1-P)HP\|_2\leq l/\sqrt{n}$ and $\lim_{n\to\infty}\|(1-P)HP\|_2=0$ (l being fixed). In order to get similar result for the cross norm stated in the theorem, we shall first prove the following formula:

$$((1-P)H\varphi_m,\ (1-P)H\varphi_k)=0, \quad \text{if} \quad m\neq k. \tag{8}$$

As $\mathfrak{M}_m$ reduces H, we get $H\varphi_m\in\mathfrak{M}_m$. But, for any $f\in\mathfrak{M}_m$, $Pf=\sum_{i=1}^{n}(f,\varphi_i)\varphi_i=(f,\varphi_m)\varphi_m\in\mathfrak{M}_m$, which implies $(1-P)f\in\mathfrak{M}_m$. Hence $(1-P)H\varphi_m\in\mathfrak{M}_m$. From this and the mutual orthogonality of $\mathfrak{M}_m$ (8) follows immediately. From (7) and (8), by making use of the Parceval's equality, we see that for any $f\in\mathfrak{H}$,

$$\begin{aligned} \|(1-P)HPf\|^2 &= \left\|\sum_{m=1}^{n}(f,\varphi_m)(1-P)H\varphi_m\right\|^2=\sum_{m=1}^{n}\|(f,\varphi_m)(1-P)H\varphi_m\|^2 \\ &\leq (l^2/n^2)\sum_{m=1}^{n}|(f,\varphi_m)|^2\leq l^2\|f\|^2/n^2, \end{aligned}$$

which means that $\|(1-P)HP\| \leq l/n$. Now we shall apply Lemma 2 and (1) to the operator $(1-P)HP$ which is clearly of at most rank n, then we obtain the estimate:

$$\alpha((1-P)HP)=\alpha([(1-P)HP]^*) \leq \varepsilon_n n \|(1-P)HP\| \leq \varepsilon_n l. \tag{9}$$

For any fixed l, we can therefore make $\alpha((1-P)HP)$ and $\alpha([(1-P)HP]^*)$ arbitrarily small by choosing sufficiently large n. Once this result is established, the rest of the proof can be performed quite similarly as in the proof of von Neumann. The only point to be noticed is the following. By virtue of (9) we can construct, as von Neumann does, a sequence of self-adjoint operators of finite rank $Y_1, Y_2, \cdots$ for which, instead of the Schmidt norm $\|Y\|_2$, $\alpha(Y_n)$ is smaller than $\varepsilon/2^n$ (ε as stated in the theorem). Then $X_n = Y_1 + \cdots + Y_n$, $n=1,2,\cdots$ form a Cauchy sequence with respect to the norm α. Hence X_n converge to an operator $X \in \mathbf{C}_\alpha$ with respect to the norm α and consequently also with respect to the ordinary norm. Thus $\alpha(X) \leq \alpha(Y_1) + \alpha(Y_2) + \cdots \leq \varepsilon$. The operator X thus obtained also satisfies the other requirement of the theorem.